

システムと操作者の対話に基づく翻訳例一般化手法の評価

秋葉 泰弘 中岩浩巳

ATR 音声言語コミュニケーション研究所
京都府相楽郡精華町光台 2-2-2
{yasuhiro.akiba, hiromi.nakaiwa}@atr.co.jp

フランシス ボンド

NTT コミュニケーション科学基礎研究所
京都府相楽郡精華町光台 2-4
bond@cslab.kecl.ntt.co.jp

1 はじめに

トランスファー方式の機械翻訳システムで用いられる意味構造変換ルール¹は、辞書作成の専門家がその条件(意味カテゴリ制約)を付与する事により、作成されている。本稿では、専門家がこなすこの作成作業を如何に支援するかと言う問題を取り上げる。実用的な機械翻訳システムを構築するためには、変換ルールを大量に作成する必要がある[8]、作成を支援する手段が必要であった。

変換ルールの作成を支援する従来法には、帰納学習によるアプローチ[7, 9]がある。このアプローチは翻訳用例を沢山必要とするため、このアプローチにより作成可能な変換ルールは限られていた。特に、テキスト中の出現頻度が低い表現(低頻度の語彙や用法)に対応する変換ルールは、大量に作成される必要である[8]にも関わらず、このアプローチでは作成が難しかった。

テキスト中の出現頻度が低い表現に対応する変換ルールの作成を支援するために、著者らはこれまで、システムがシステムと専門家のやり取りを通して変換ルールの条件を獲得する支援手法を提案してきた[2, 3]。以下、同システムを汎化支援システム、または単に、支援システムと呼ぶ。支援システムは、シソーラスを利用して例文を生成し、その例文が翻訳例として正しいか否かを専門家に質問し、専門家の応答により適切な条件を探索する。支援システムが提示する質問に回答するために必要な能力は、原言語を目的言語に翻訳する能力だけである。従って、システム操作は辞書作成の専門家でなくとも可能である。

本稿では、支援システム[2]を操作する人が辞書作成の専門家である場合と、翻訳の知識のみを有する人(非専門家)である場合とで、作成される変換ルールにどの程度差があるかを評価した結果について報告する。実験では、各群二名に支援システムを操作して貰い、100個の意味カテゴリを同定する作業を行った。その結果、作成される意味構造変換ルールには殆んど大差がない事がわかった。

¹日本語語彙大系[6]では構文体系と呼ばれている。

以下、2節では本稿で取り上げる支援システムの概要を説明する。3節で実験の方法及び結果を示し、その結果を議論する。4節でまとめる。

2 汎化支援システムの概要

2.1 獲得タスク

トランスファー方式の機械翻訳システムでは、動詞と名詞の共起関係に着目し、原言語の共起関係パターンとそれに対応する目的言語の共起関係パターンの対からなる意味構造変換ルール(以下、パターン対と呼ぶ)が中心的な役割を果たす。例えば、NTTで研究開発している日英機械翻訳システム、ALT-J/E[5]は、このようなパターン対を備えた、トランスファー方式の機械翻訳システムである。

図1に ALT-J/Eのパターン対の例を示す。IF部は日本語パターンで、THEN部はそれに対応する英語パターンである。《人》、《文書》、《心》が入っている各スロットには、意味カテゴリと呼ばれる名詞の語義²が入る。ALT-J/Eは、意味カテゴリを約2700個持っており、これらの意味カテゴリは最大の深さ13段の階層構造(シソーラスと呼ぶ)を成す。ALT-J/Eの日本語辞書中の各名詞(全部で約37万語)は、いずれかの意味カテゴリ(1つとは限らない)を語義に持つ。パターン対中の意味カテゴリは、形態素解析や構文解析にも用いられ、意味的制約を用いずには解決できないような、解析多義を解消するためにも利用されている。辞書作成の専門家は、パターン対を作成

²日本語語彙大系[6]では、上位概念の《抽象(精神)》及び《精神》が指定されている。

IF	J-Verb	= '読む'
	N ₁ (J-Subj)	≡ 《人》
	N ₂ (J-Obj)	≡ 《文書》 or 《心》
THEN	E-Subj	= N ₁
	E-Verb	= 'read'
	E-Obj	= N ₂

図1: パターン対の例(日本語動詞「読む」に対する)

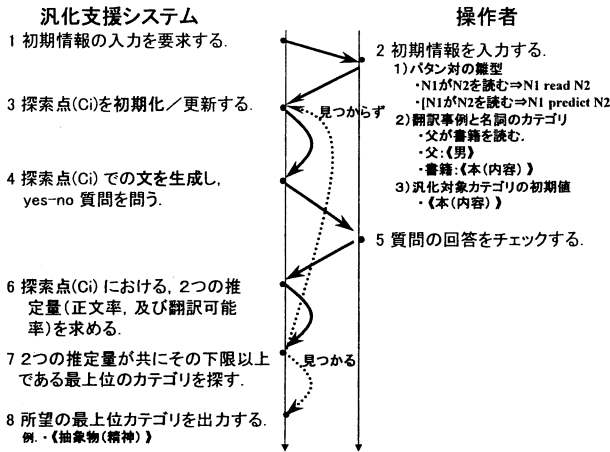


図 3: 汎化支援システムの概要

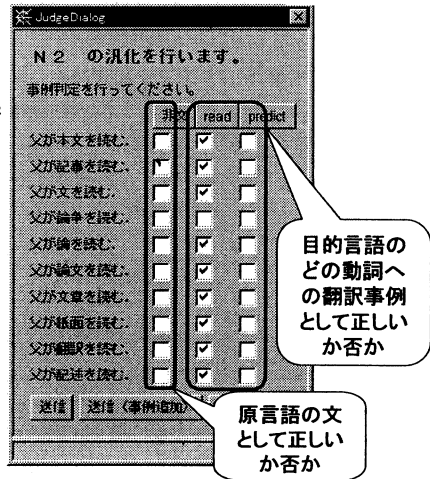


図 4: 質問回答画面の例

[/ 〈名詞〉 / 〈抽象〉 / 〈抽象物〉]

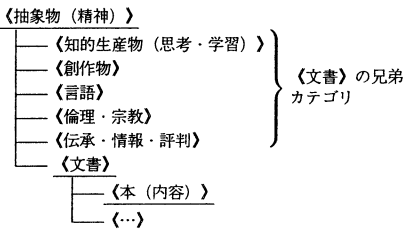


図 2: 日本語語彙大系，一般名詞意味属性体系の一部 [6]

する際、各スロットに入れるべき適切な（抽象的過ぎず、特殊過ぎない）意味カテゴリを探索し、指定する。著者らは、作成作業を支援するために、この探索をシステム化した [2, 3]。

2.2 汎化支援システム

本節では、本稿で取り上げる支援システム [2, 3] の概要及びポイントについて説明する。支援システムは、図 3 に示すように、翻訳事例を支援システムとその操作者の対話（質問応答）に基づいて一般化（汎化）を行い、指定する意味カテゴリを探索する³。

支援システムはまず、システム操作者から、ボタン対の雛形⁴（例えば「N1 が N2 を読む ⇒ N1 read N2」）、翻訳事例（例えば「父が書籍を読む」）、事例中の名詞の意味カテゴリ（例えば「本」に対して《男》、「書籍」に対して《本（内容）》）、汎化対象の意味カテゴリの初期値（例

³ボタン対を完成させるには、指定する意味カテゴリ毎に探索を繰り返す。

⁴雛形作成法の研究としては、[4]がある。

例えば《本（内容）》）を入力して貰う。

取り上げている例では、N2のスロットに入る意味カテゴリを探索する場合支援システムは、《名詞》と《本（内容）》を結んだパス上の意味カテゴリ（図 2 で、下線が引かれた意味カテゴリ）から適切な意味カテゴリを探索する。パス上の意味カテゴリ（例えば《抽象物（精神）》）が直下の意味カテゴリ（《文書》）の代りに指定された場合、直下の意味カテゴリの兄弟カテゴリ（《知的生産物（思考・学習）》～《伝承・情報・評判》の 5 カテゴリ）またはこれらの子孫カテゴリを語義に持つ名詞がボタン対にマッチするようになる。そのような全名詞のうち、1) どのくらいの割合が「書籍」と置き換わっても原言語として正しい文を生成され得るか、2) どのくらいの割合が置き換わった文が目的言語の動詞の翻訳事例となり得るか、により、より抽象度の高い意味カテゴリを指定すべきか否かを判定し、探索を繰り返す（詳細は、[2, 3]を参照）。

1) や 2) の情報を得る為に、支援システムは、名詞の置き換えにより生成した文を図 4 に示すようにシステム操作者に提示し、システム操作者に判断して貰う。

置き換えを行う名詞をシステムが全て提示し、操作者に判断して貰うのは非現実的であるため、置き換えを行う名詞は前述の兄弟カテゴリまたはそれらの配下の名詞のうち、一部に限る必要がある。[3]では、各兄弟カテゴリ配下のカテゴリからランダムに選択した名詞で置き換えを行う。[2]では、前述の割合の推定精度を向上させるために、各兄弟カテゴリまたはその子孫カテゴリ配下のカテゴリからその最下位概念数の比づつ（図 2）、名詞を使用頻度の高い順に選び出す。各意味カテゴリ配下の名詞の使用

頻度は、[1]により予めコーパスから計算しておく。

名詞の置き換えによって生成された文は、システム操作者により、1) 提示された文が原言語として正しいか、2) 提示された文が目的言語の動詞の翻訳事例となり得るか、を判断される。(具体的な判断事例は、[2]を参照。)この判断は翻訳に知識がありさえすれば可能であり、必ずしも、辞書作成の専門家である必要はない。

3 実験

支援システム[2]を操作する人が辞書作成の専門家(A群)である場合と、翻訳の知識のみを有する人(非専門家, B群)である場合とで、1) 作成される変換ルールにどのような差があるか?, 2) 作成に要する時間にどの程度差があるか?, を調べる実験⁵を行った。

3.1 実験条件

実験では、日英機械翻訳システム ALT-J/E向けに入手で既に作成した、ボタン対49個⁶に対して指定されている意味カテゴリ100個を支援システムを利用して同定する作業を行った。これらのボタン対は、使用頻度が少ない日本語動詞や日本語動詞の用法に対応するわけではないが、指定されている意味カテゴリが妥当であることがチェックされている必要があったので、これらを用いた。以下、同定する意味カテゴリを想定カテゴリと呼ぶ。

A群の操作者はボタン対作業経験者2名⁷とし、B群の操作者はボタン対作業経験者はないが、英語の翻訳能力はある2名とした。実験に際しては、各群2名の作業者に個別に支援システムを操作して貰い、支援システムが同定する意味カテゴリを調べた。支援システムが行う探索は、トップダウン探索、ボトムアップ探索の何れかでを行った。支援システムに入力する初期情報(ボタン対の雛形、翻訳事例、名詞の意味カテゴリ)は実験の前に辞書作成の専門家に選んで貰い、各システム操作者の開始条件⁸を揃えた。汎化対象カテゴリの初期値としては、ボタン対に指定するカテゴリを大雑把に支援システムに与える事を想定しているため、トップダウン探索の場合は想定カテゴリの一段上のカテゴリを入力し、ボトムアップ探索の場合は一段下のカテゴリを入力した。また、実験で利用したシソーラ

⁵本稿では、獲得された意味カテゴリの良し悪しを直接調べたかったため、機械翻訳システムに組み込んで翻訳の品質がどう変わるかは調査の対象としなかった。

⁶意味カテゴリの指定が3個程度(ボタン対の平均指定数)で、訳分けの数が著しく多くないボタン対を選んだ。

⁷ボタン対作成経験者はこの2名のみであるため、各群の作業人数は2名とした。

⁸意味カテゴリの良さを測る2つの割合(正文率及び翻訳可能率)の下限は、順に、2%と80%とした。

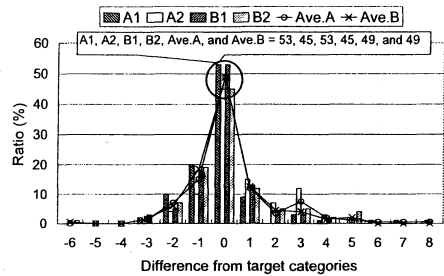


図5: トップダウン探索で支援システムが獲得したカテゴリの位置(想定カテゴリに対する相対位置)の、全試行における割合

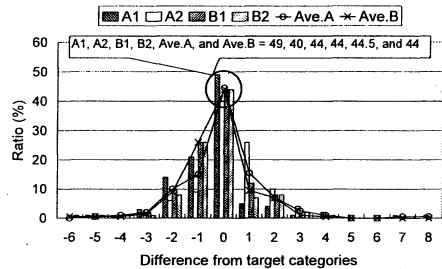


図6: ボトムアップ探索の場合の、図5対応する結果

は ALT-J/E で採用されている、日本語語彙大系の一般名詞意味属性体系である。

3.2 実験結果

図5, 6は、順に、トップダウン探索及びボトムアップ探索で支援システムが獲得した意味カテゴリの位置(想定カテゴリに対する相対位置)の、全試行における割合を示す。例えば、'-1'に対する棒グラフは、A群2名(A1, A2)、B群2名(B1, B2)各々が支援システムを操作した際、獲得した意味カテゴリが想定カテゴリの1つ下である試行が全100試行中どの位の割合であったかを示す。Ave.A及びAve.Bは、各群内の平均を示す⁹。群内の差はあるが、平均的にはA群とB群で、得られる結果に大きな差は無かった。なお、B群2名の結果が一致した割合は、トップダウン探索、ボトムアップ探索の順に、28%、32%(約3割)で、その結果が想定カテゴリと一致する割合は、53%、62%。B群2名何れかの結果が想定カテゴリである割合は、73%、65%であった。B群2名何れかの結果も想定カテゴリでない場合の大半は、想定カテゴリより下位(1, 2段下)の意味カテゴリが得られた。厳密には得られた意味カテゴリの方が正しいが、厳密解を取る場合、兄弟カテゴリやその子孫カテゴリの一部な

⁹相対位置の平均は、両群共に1段以内(約0.9段)であった。

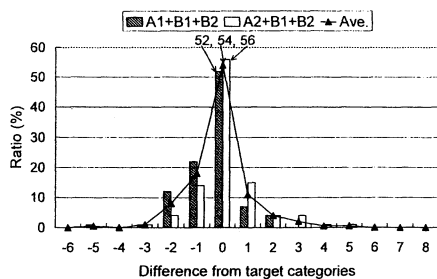


図 7: 3名の操作者各々がトップダウン探索により獲得した意味カテゴリのメジアン位置(想定カテゴリに対する相対位置)の、全試行における割合

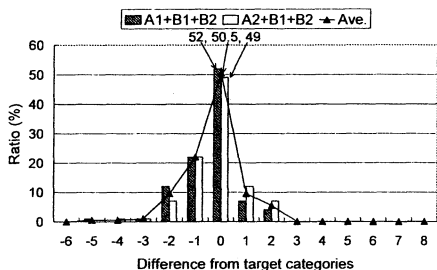


図 8: ボトムアップ探索の場合の、図 7に対応する結果

ど、複数の意味カテゴリを同時に指定する必要が生ずるため、専門家は想定カテゴリを採用していた。

図 7, 8は、順に、トップダウン探索及びボトムアップ探索で3名の操作者各々が支援システムを操作して獲得した意味カテゴリのメジアン位置(想定カテゴリに対する相対位置)の、全試行における割合を示す。B群内のメジアンを導出する為に、B群2名とA群1名¹⁰でメジアンを取った。A2の結果はB群2名より劣るが、B群2名とA2の結果のメジアンが想定カテゴリに一致する割合は、個々の結果が想定カテゴリに一致する割合より3%~11%向上する。複数名の結果のメジアンを候補とすると、個々の性能より向上する傾向が見られる。

表 1に操作時間を示す。操作時間は、支援システムに初期情報を入力する時間、及び質問応答を繰り返し、1つの意味カテゴリを得るまでに掛った時間を測った。ボタン対を獲得するための合計時間は、「前者の時間と後者の時間の和の、付与する意味カテゴリ数倍」で概算できる。[6]では平均3個の意味カテゴリが指定されている。B群の操作者が、A群に比べ操作時間が大きく長くなる事は無かった。この支援システムを利用して1つのボタン対を作成するのに掛る時間は、概算で13分で、この時間は辞書

¹⁰平均的には、A群とB群は大きな差がないので、1名はA群の1名の結果で代用した。

表 1: 支援システム操作時間 (X:YはX分Y秒を表す.)

トップダウン探索							
初期設定入力時間				1カテゴリ汎化所要時間			
A1	A2	B1	B2	A1	A2	B1	B2
1:27	2:33	1:34	1:38	2:21	2:39	2:09	3:07
2:00 (Ave.A)		1:36 (Ave.B)		2:30 (Ave.A)		2:38 (Ave.B)	
1:48 (Total Ave.)				2:34 (Total Ave.)			
ボトムアップ探索							
初期設定入力時間				1カテゴリ汎化所要時間			
A1	A2	B1	B2	A1	A2	B1	B2
1:21	2:23	1:37	1:38	1:36	2:42	2:18	3:20
1:52 (Ave.A)		1:37 (Ave.B)		2:09 (Ave.A)		2:49 (Ave.B)	
1:45 (Total Ave.)				2:29 (Total Ave.)			

作成の専門家が人手で作成する時間の約1/7である。

4 おわりに

本稿では、トランスファー方式の機械翻訳システムで用いられる意味構造変換ルールを作成する支援システムを概説した。支援システム [2] を操作する人が辞書作成の専門家である場合と、翻訳の知識のみを有する人(非専門家)である場合とで、作成される変換ルールにどのような差があるかを実験評価したところ、非専門家でも辞書作成の専門家が支援システムを使って得られるのとはほぼ同等の意味カテゴリが付与できた。また、非専門家2人の操作結果を合わせると、辞書作成の専門家が人手で付与するであろう意味制約が約7割得られた。

参考文献

- [1] 秋葉泰弘, 中岩浩巳. コーパスを用いた意味カテゴリの代表的語彙の収集法. 第61回情報処理学会全国大会講演論文集, 第2巻, pp. 137-138, 10/3-10/5 2000.
- [2] Yasuhiro Akiba, Hiromi Nakaiwa, Yoshifumi Ooyama, and Satoshi Shirai. Interactive generalization of a translation example using queries based on a semantic hierarchy. *International Journal of Artificial Intelligence Tools (IJAIT)*, Vol. 10, No. 4, pp. 675-690, 12 2001.
- [3] 秋葉泰弘, 白井諭. 原言語文とその対訳に関する質問に基づく意味構造変換ルールの獲得. 言語処理学会第5回年次大会発表論文集, pp. 430-433, 3/16-3/18 1999.
- [4] Sanae Fujita and Francis Bond. Extending the coverage of a valency dictionary. In *Proc. of COLING-2002 workshop on Machine Translation in Asia*, pp. 67-73, Taipei, 2002.
- [5] 池原悟, 宮崎正弘, 白井諭, 林良彦. 言語における話者の認識と多段翻訳方式. 情報処理学会論文誌, Vol. 28, No. 12, pp. 1269-1279, 1987.
- [6] 池原悟, 宮崎正弘, 白井諭, 横尾昭男, 中岩浩巳, 小倉健太郎, 大山芳史, 林良彦. 日本語語彙大系. 岩波書店, 東京, 1997.
- [7] Hang Lie and Naoki Abe. Generalizing case frames using a thesaurus and the MDL principle. *Computational Linguistics*, Vol. 24, pp. 71-88, 1998.
- [8] 白井諭, 池原悟, 横尾昭男, 井上浩子. 日英機械翻訳に必要な結合価ボタン対の数とその収集方法. 情処研報 NL-110-7, pp. 43-50, 1995.
- [9] Takehiro Usturo. Sense classification of verbal polysemy based on bilingual class/class association. In *Proc. of COLING-96*, pp. 968-973, 1996.